Oecologia Australis
21(1): 42-53, 2017
10.4257/oeco.2017.2101.04

USING DIFFERENT PROXIES TO PREDICT HANTAVIRUS DISEASE RISK IN SÃO PAULO STATE, BRAZIL

Paula Ribeiro Prist[1], Renata de Lara Muylaert[2], Amanda Prado[1], Fabiana Umetsu[4], Milton Cezar Ribeiro[2], Renata Pardini[3] & Jean Paul Metzger[1]*

[1] *Universidade de São Paulo (USP), Instituto de Biociências, Departamento de Ecologia, Rua do Matão, 321, travessa 14, São Paulo, SP, Brasil, CEP: 05508-900*
[2] *Universidade Estadual de São Paulo (UNESP), Instituto de Biociências, Departamento de Ecologia, Laboratório de Ecologia Espacial e Conservação (LEEC), Av. 24ª, Bela Vista, Rio Claro, SP, Brasil, CEP: 13506-900*
[3] *Universidade de São Paulo (USP), Instituto de Biociências, Departamento de Zoologia, Rua do Matão, 321, travessa 14, São Paulo, SP, Brasil, CEP: 05508-900*
[4] *Instituto Federal de Educação Ciência e Tecnologia de Farroupilha (IFRS), Campus Alegrete, RS-377, Km 27, Passo Novo, Alegrete, RS, Brasil, CEP 97555-000*
Emails: pprist@hotmail.com, renatamuy@gmail.com, amandaprado07@gmail.com, fabiume@yahoo.com; mcr@rc.unesp.br, pardinirenata@gmail.com; jpm@ib.usp.br*

ABSTRACT
Recent studies predict disease risk using different proxies, such as pathogen prevalence in hosts, abundance of the main hosts, and the number of reported disease cases. These proxies are used to build risk maps that can aid the prevention of new disease outbreaks. To date, these proxies have not been widely tested for differences in their predictions and effectiveness, which could have serious implications for disease control measures. In this study, we compared two different proxies inferring hantavirus disease risk in the state of São Paulo. We compared risk level distribution to the accuracy of risk maps using (a) Rodent Reservoir Abundance data (RRA) sampled in 2002-2008 and (b) Hantavirus Pulmonary Syndrome cases reported (RC) in 1993-2012. RRA data were collected within forest fragments and in the matrix of six landscapes, and were extrapolated for São Paulo State through regression models using the amount of forest cover and the collection context as predictors. Using Bayesian models, we created a HPS risk map using annual HPS incidence, climate, landscape structure metrics and social factors. We validated RRA and RC risk maps with actual reported HPS cases (2013-2015). These data were categorized according to risk levels and compared using histograms and correlations. The two risk maps (RRA and RC) had a low Pearson correlation (0,038) and a low covariance (0,016), indicating high uncertainty in the predictions between these two proxies. The RRA map predicted that 68% of the municipalities in the state are in the medium to high risk categories, while the RC map predicted only 6%. This indicates that the RRA risk map might be overestimating high risk areas. The RRA map also had a higher sensitivity than the RC map to newly reported cases, correctly identifying 82% of the cases in medium to high risk areas. On the other hand, the RC map had a higher specificity (91%), leading to better prediction of the low risk areas (31% for RRA map). Our results draw attention to the fact that different proxies can give different results and predict different risk levels and should be used carefully in disease studies.

Keywords: emergent diseases; landscape epidemiology; predictive power; risk maps; rodent host abundance.

INTRODUCTION

Some studies use different proxies to infer transmissible disease risk via maps, which have become more common in the last 20 years (Kitron 1998). The proxies generally use presence/abundance of the main hosts in the environment (Guerra *et al*. 2002). Pathogen prevalence in hosts (Ostfeld *et al*. 2005, Xiao *et al*. 2016) and the number of disease reported cases (Bhatt *et al*. 2013) can also be used. This information, when coupled with information on host ecological requirements, enables the prediction of patterns of disease emergence, spread and control (Biek & Real 2010), using different methods. One of these methods is mapping host distribution to infer disease risk, as used for Lyme disease (Guerra *et al*. 2002), malaria (Martens *et al*. 1999), and dengue (Little *et al*. 2011). Likewise, the number of infected hosts or reported disease cases can also be used in Bayesian models, regression or niche modeling analysis, as used for hantavirus pulmonary syndrome (HPS) (Rogers & Randolph 2000, Glavanakov *et al*.

2001) and bluetongue in France (Guis *et al*. 2007).

Hosts biology dictates the dynamics of pathogen transmission and persistence (Reisen 2010). For example, Lyme disease and HPS hosts may serve as effective long-term reservoirs (Tälleklint *et al*. 1995), with the effective dispersal of the pathogen being limited to the dimensions of the reservoir home ranges (Madhav *et al.* 2004). However, the presence or high abundance of hosts alone does not guarantee that humans will become infected. To acquire a zoonotic disease, human exposure to infected hosts is also necessary; disease transmission results from a combination of human risk behavior and host risk. Human risk behavior includes proximity to host habitat due to anthropogenic activities, while host risks include host density and the proportion of infected hosts (Horobik *et al*. 2007). Anthropogenic influences affect the epidemiology of several diseases and transmission dynamics by altering the landscape structure, providing mechanisms that enhance host breeding success and increase disease risk.

Disease risk assessment comes with several caveats (Ostfeld *et al.* 2005), particularly because it is hard to incorporate all important risk factors in a single model. The use of spatially referenced and temporal data may add strength to epidemiological analyses. However, to date, no study has evaluated the use of different data (host presence, infection prevalence, disease cases) to infer disease risk, nor integrated this information with data derived from remote sensing.

HPS ranks among the major emerging diseases of the last century, with great potential to become a public health threat in the near future (Pereira *et al*. 2007). It was first recognized in 1993, in USA (CDC 2014) and Brazil (Brazilian Ministry of Health 2013). Rodents of the family Cricetidae are the primary hosts of hantavirus (Jonsson *et al*. 2010, family *Bunyaviridae*), a virus that causes two syndromes in humans: HPS in the Americas and hemorrhagic fever with renal syndrome in Eurasia and Africa (Jonsson *et al*. 2010). Transmission to humans occurs via inhalation of aerosolized virus particles emitted from the excreta of infected rodents (Lee *et al.* 1981, Vapalahti *et al.* 2010). HPS has a high lethality rate, especially in Brazil, where it reaches ~40% (Brazilian Ministry of Health 2013, de Oliveira *et al.* 2015).

Each Brazilian region has different reservoir species that host distinct hantavirus strains (Guterres *et al.* 2015). In southern Brazil, the sigmodontine rodents *Oligoryzomys nigripes* (Olfers, 1818) and *Necromys lasiurus* (Lund, 1840) are the main reservoirs of human HPS from the Juquitiba and Araraquara viruses, respectively (de Oliveira *et al.* 2013). Both species are adapted to inhabit anthropic environments (Umetsu & Pardini 2007, Gheler-Costa *et al.* 2012) and native habitat edges (de Oliveira *et al*. 2015).

Evaluating the suitability of different data sources for predicting disease risk to ascertain the best one can result in more accurate predictions of future outbreaks. Better inferences of risk are critical to define effective surveillance and control programs, with the need for studies that include host data in disease risk analysis (de Oliveira *et al.* 2015). However, the different outcomes and efficacies of disease risk prediction by different proxies (abundance data and disease cases) remains unexamined. Here, we fill this research gap by using rodent reservoir abundance (RRA) data and the number of reported HPS cases (RC) in a case study focused on São Paulo State, Brazil. We created risk maps from both datasets and hypothesized that the similarities between the two risk maps will be intermediate, since high abundances of reservoir rodents may lead to increased contact rates and increased rates of infection for both rodents and humans, increasing HPS transmission risk; and that the risk map based on HPS reported cases will have more accuracy in predicting high risk areas than the risk map based on rodent reservoir abundance data, since contracting diseases requires not only a certain threshold of reservoir abundance but also human risk behavior, which is included in the data used to create this risk map.

MATERIAL AND METHODS

Study area

In this study we used data from São Paulo State, Brazil, which includes both the cerrado and the Atlantic Forest biomes and cover an area of 248,200 km^2. This region is ideal for this type of study because it presents

a marked gradient of forest cover (Ribeiro *et al.* 2009), different crops and human population sizes, all important factors of disease risk.

Rodent abundance data

Rodents were sampled between 2002 and 2008, using a 100-m sequence of 11 pitfalls traps (60 L) at 104 sites across six Atlantic Forest landscapes in São Paulo State. Sampling regions included three fragmented landscapes (11%, 31% and 49% of native forest cover; here called FClandscapes) and three continuous forest (>90% native forest cover) (Pardini *et al.* 2010, Umetsu 2010; Figure 1). From the 104 sites, 68 were in native forest: 50 in fragments (15 fragments in 11% native forest cover; 15 in 49% and 20 in 31%), and the other 18 in continuous forest (for more details see Pardini *et al.* 2010). The remaining 36 sites were in annual crops (landscapes with 11% and 49% native forest cover) (Umetsu 2010). The anthropogenic matrix of the landscapes included crops, pastures, urban areas, forestry and native forest in early successional stage. For continuous and fragment landscapes, capture sessions were conducted during 2002 and 2003, totaling 23,936 trap-nights (Pardini *et al.* 2010). For agricultural sites, capture sessions were conducted during 2008, totaling 6,336 trap-nights (Umetsu 2010).

Oligoryzomys nigripes was collected in matrix, fragments and continuous forest (here called collection context). For matrix sites, we calculated the abundance per site as the sum of all individuals captured over all capture sessions. For animals collected in forests, we calculated the abundance per site as the sum of all individuals captured over all capture sessions divided by two, because the capture effort at forest sites was double that for matrix sites. For *N. lasiurus* we consider only the data collected at the 36 matrix sites, because the number of individuals collected in forest landscapes was very small. The captured rodents were not tested for infection rates; therefore, given that we did not have this information for a gradient of environmental conditions, we used only the host abundances. Rodents were always collected in the summer and in the same period as the epidemiologic data (1993-2012), which corresponds to the collection periods of our predictor variables. We compared the abundances in our RRA model with the results of other studies conducted in São Paulo State, using abundance standardized by the sampling effort for *O. nigripes* and *N. lasiurus* separately.

Disease and social data

Data on HPS incidences per municipality per year for 1993-2012 were obtained from the Center for Epidemiological Surveillance of the São Paulo State. Thus, we considered the 645 municipalities of São Paulo State as our sampling units. For each municipality we assigned a binary variable according to the presence (value 1) or absence (value 0) of HPS.

Within rural landscapes, HPS generally occurs in men over the age of 20, who live or work in agricultural areas (de Oliveira *et al.* 2014, Willemann & de Oliveira 2014). This may occur mainly because of the absence of preventive measures (Ferreira 2003). Therefore, we used the number of rural men older than 14 years in each municipality, obtained from the National Institute of Geography and Statistics (IBGE; www.ibge.gov.br), as the population at risk for HPS, and the Human Development Index (HDI) served as a proxy for human development. This information was available only for 1996 and 2006, but we wanted to model the incidence of HPS from 1993 to 2012. Therefore, we used the 1996 data as covariates to predict disease incidence for 1993-2001 and the 2006 data to predict incidence for 2002-2012. We collected municipal HDI data for 1991, 2000 and 2010 from IBGE, using the 1991 data to predict incidence for 1993-1998, the 2000 data to predict incidence for 1999-2005, and the 2010 data to predict incidence for 2006-2012.

Landscape metrics for modelling HPS

Metrics for reported cases (RC) map

We used the São Paulo State forest maps (www.iflorestal.sp.gov.br) for 2000 and 2010 and the programs ArcGis 10.0 and Fragstats 4.1 (McGarigal *et al.* 2002) to calculated landscape metrics (percentage of native habitat cover and number of habitat fragments) for the native vegetation of each municipality in São Paulo State,

aggregating both forest and cerrado vegetation patches. To match this information with available disease data, we used metrics extracted from the 2000 map as covariates to model incidence for 1993-2001, and metrics extracted from the 2010 map served as covariates for the period 2002-2012. The proportion of sugarcane cultivated in each municipality was obtained from the Institute of Agricultural Economics (www.iea.sp.gov.br). For the analyses presented here, we used annual agricultural data (1993-2012) to predict annual disease incidence from 1993-2012.

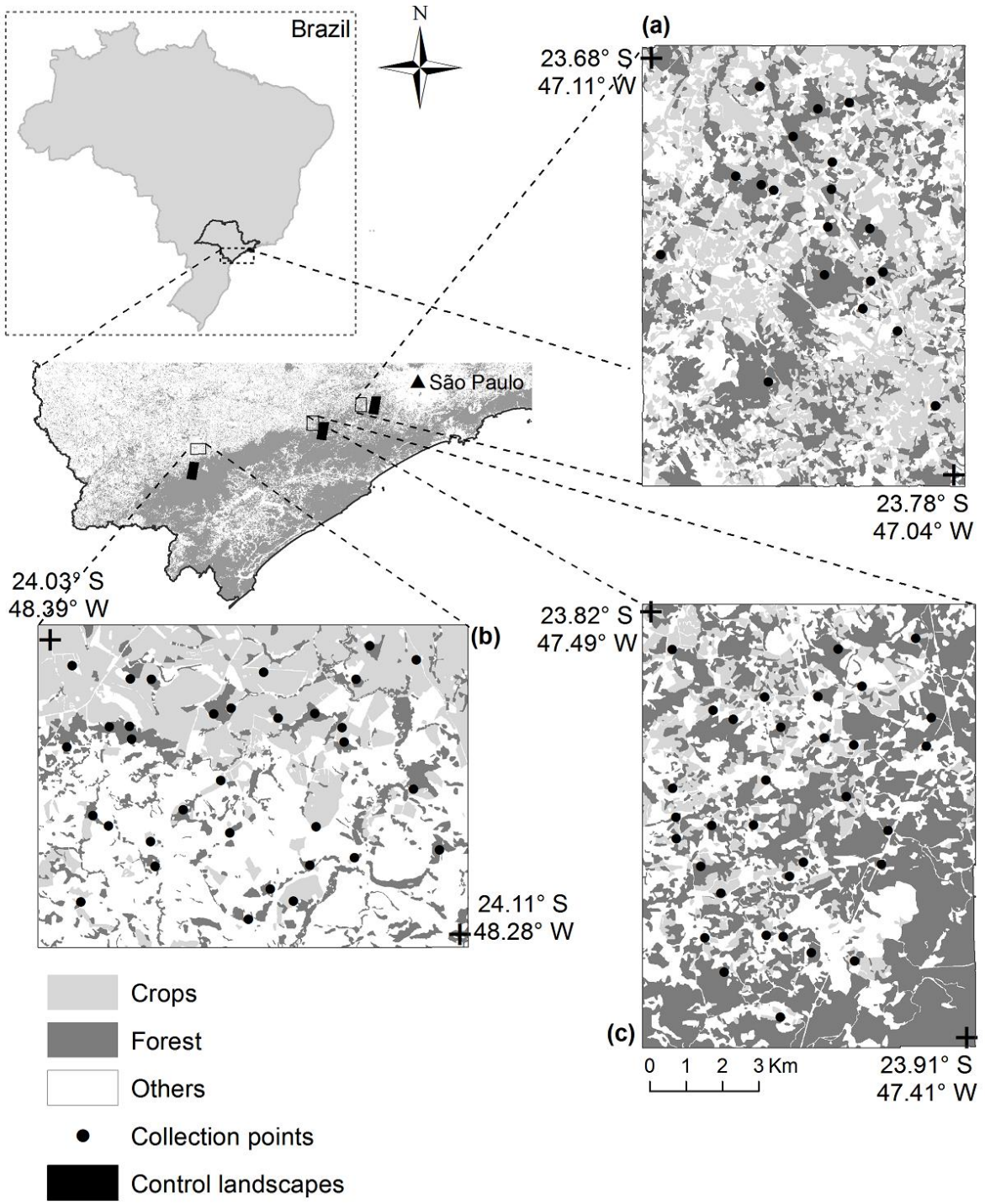

Figure 1. Location of the studied landscapes in São Paulo State, where rodent abundance data were collected. (a) 31% forest cover landscape, (b) 11% forest cover landscape, and (c) 49% forest cover landscape. Black dots show collection points and the black rectangles are the control landscapes of continuous forest sites. Adapted from Pardini *et al.* (2010).

Metrics for rodent reservoir abundance (RRA) map

We used the São Paulo State forest map for 2010, in ArcGis 10.0, to calculate landscape metrics (percentage of native habitat cover and the amount of edge density) for the native vegetation cover of each scale: 200-m, 500-m and 800-m radius circles around each rodent collection point. These scales were set according to the species' rate of movement (50 to 100 m) and the study of Jackson and Fahrig (2012), which suggests that the appropriate extent is between four to nine times the average dispersal distances of a species of interest.

Climatic variables

Meteorological data was obtained from the International Research Institute for Climate and Society (http://iridl.ldeo.columbia.edu/index.html). NOAA/ NCEP provided temperature data, at ~50-km spatial resolution and monthly time step; and CHIRPS provided precipitation data, at ~5-km spatial resolution and 10-day precipitation average monthly step. We calculated annual mean temperatures and total annual amounts of precipitation for 1993-2012 for each municipality and used this data to predict annual disease incidence.

Statistical analysis

Predicting HPS using RRA

Using a model selection approach, we fitted generalized linear models with the following variables: percentage of native habitat cover and the amount of edge density at each scale; the collection context; the percentage of native forest cover at landscape level for *O. nigripes;* and the percentage of native forest cover (FC) at each scale, percentage of native forest cover at landscape level, and type of crops ('corn' or 'others') for *N. lasiurus*. All predictor variables included in the models had a correlation lower than 21%. The best models selected for each species were those with the lowest AICc values (FC at 800-m scale, FClandscape, and the interaction between collection context and FClandscape for *O. nigripes*; and FC at 800-m scale for *N. lasiurus*; Table 1). From the best models, we extrapolated the abundance of the two species to the entire state, using the São Paulo State forest map for 2010. The *N. lasiurus* and *O. nigripes* maps were summed to build a unique abundance map for both rodents using ArcGIS 10.0.

Predicting HPS using RC

The probability of hantavirus infection risk for the state of São Paulo was calculated as a function of landscape, social and climatic factors, using a Bayesian model described in detail in Prist *et al*. (2016). HPS infection risk was predicted using a Bernoulli distribution and a model containing seven fixed covariates: proportion of sugarcane; proportion of native vegetation cover; number of native vegetation patches; HDI; mean annual temperature (°C); total annual precipitation (mm); and rural male population >14 years old. Municipality was included as a random effect to account for differences among them. All estimated parameters were standardized by centering them on their mean and dividing by two standard deviations (Gelman & Parode 2006). All priors were assigned as uninformative distributions, and model convergence and performance was examined via Gelman-Rubin diagnostics. Model results were used to generate a map of hantavirus risk areas for the state of São Paulo.

Table 1. Slope and standard error (±SE) for each variable from the best models explaining the abundance of *Oligoryzomys nigripes and Necromys lasiurus* (RRA model).

Predictor Variables	***O. nigripes***	***N. lasiurus***
% native habitat cover at 800m	0.008 (±0.004)	-0.033 (±0.011)
FCLandscape	-0.00019 (±0.007)	
Context (Fragments)	2.472 (±0.331)	
Context (Forest)	-7.273 (±17.825)	
Context (Fragments): FClandscape	-0.043 (±0.008)	
Context (Forest): FClandscape	0.081 (±0.198)	

Risk map comparison and validation

The HPS risk map based on reported cases (RC) was classified as small (<5%), medium ($5 and #10%), high ($10 and #20%) and extremely high (#20%) risks. The HPS risk map based on rodent reservoir abundance (RRA) was classified as small (<30%), medium (30$ and #40%), high (40$ and # 45%) and extremely high (>45%) risks, allowing comparison. Classifications were done in ArcGIS using break classes according to quantities based on natural groupings inherent in the data.

Risk maps were then compared through histograms illustrating the differences in their distributions for each risk category, and through correlation (*i.e.*, indicating the relationship between two datasets; calculated as the ratio of the covariance between the two layers, divided by the product of their standard deviation) and variance (*i.e.*, statistical measure of variance from the mean). We did this using the Band Collection Statistics tool in Spatial Analyst-Multivariate (Batista *et al.* 2016). Continuous data for both risk maps were compared through Pearson's correlation. Thus, we ascertained whether the maps were comparable in predicting the levels of risk in the same municipalities. Subsequently, risk maps were validated with actual reported HPS cases (2013-2015) to see which proxy had the best sensitivity (*e.g.*, proportion of correctly identified positives) and specificity (*e.g.*, proportion of correctly identified negatives) (Brooker *et al.* 2002) in predicting new disease cases, which we considered as indicators of accuracy. We measured sensitivity as the percentage of municipalities with actual disease cases in higher risk categories (medium, high and extremely high), and specificity was the percentage of municipalities without actual infection and/or disease that were correctly categorized as lower (small) risk.

RESULTS

The two risk maps, one obtained from rodent reservoir abundance data and the other from reported cases, had a low correlation in terms of risk categories, with r=0.038 of accordance and 0.017 of covariance. Considering risk data without categorization, correlation between RC and RRA maps also showed a low correlation: (r=0.11, p-value = 0.003). The results from the statistical model generated through rodent abundance data classified 208 municipalities as low risk (32% of the state) and 437 as medium or high risk (68%) for HPS infection (Figure 2). According to the RRA map, the highest infection risk is present in municipalities in the west and north region, while the eastern part of the state, which includes the Serra do Mar, is classified as low risk for HPS infection (Figure 3).

The results from the model generated through reported cases classified 6% of the state in the medium (5-10%) or high (>10%) risk category for HPS infection, and 94% was classified in the low risk (<5%) category. These numbers means that 606 (94%) out 645 municipalities were classified as low risk, 21 (3.2%) as medium risk, 13 (2%) as high risk and 5 (0.8%) as very high risk (Table 2; Figure 2). According to this risk map, the highest infection risk is present in municipalities in the northeast region, followed by some municipalities in the east, close to Serra do Mar, and in the western part of the state (Figure 3).

From 2013 to 2015, 33 new HPS cases were recorded in the state of São Paulo. The HPS risk map based on reported cases hit 36% (=sensitivity) of the new cases, which were reported in municipalities classified as medium to extremely high disease risk. The other 64% of cases (21) were reported in municipalities with low risk for HPS. From these 21 cases, 15 were reported in municipalities with up to 2% risk. From the 606 municipalities classified as low risk, only 19 had HPS cases during these two years - an error of 3.13%. Additionally, RC risk map predicted that 587 (91%) municipalities would be at low risk for HPS infection, and these presented zero new cases of the disease.

The RRA map identified 82% (=high sensitivity) of the new cases, which were reported in municipalities classified as medium to high risk. From the 208 municipalities classified as low risk, only six had actual HPS cases - an error of 2.88%. Furthermore, the RRA map predicted that 202 municipalities (31%) would be at low risk for HPS infection. Both risk maps classified municipalities in the eastern part of the state, close to the Serra do Mar region, in the low risk category, and some of the municipalities in the northwest region were categorized as high risk.

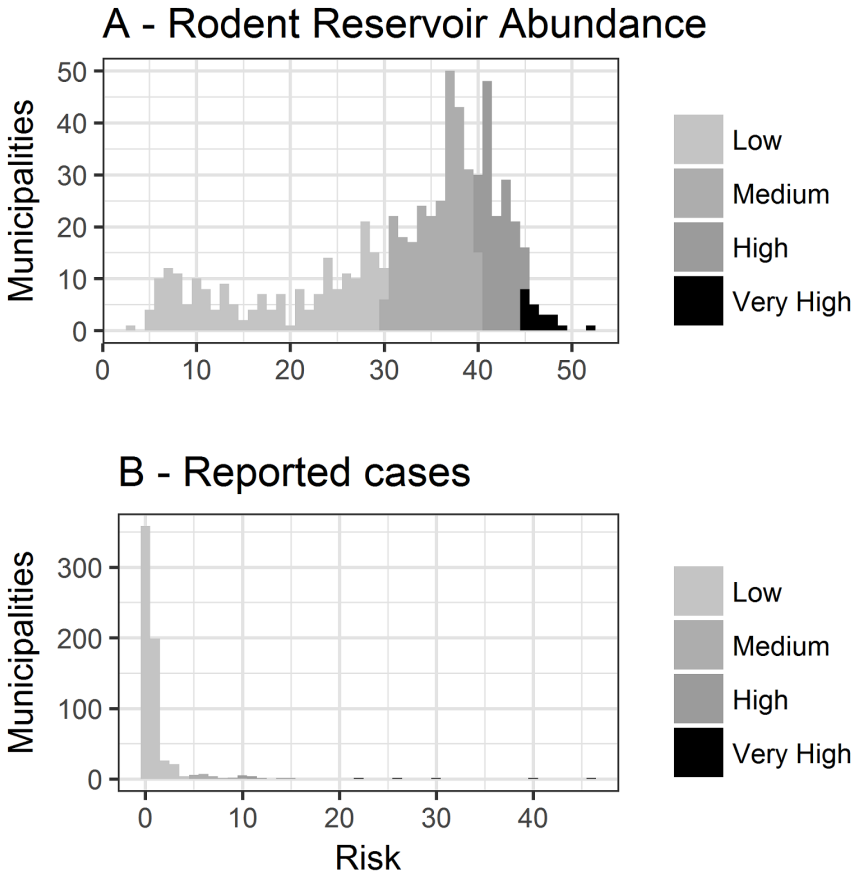

Figure 2. Number of municipalities classified in each risk category (low, medium, high and very high) according to (a) the reported cases risk map and (b) rodent reservoir abundance.

Table 2. Number of municipalities classified in each risk category of each risk map, predicted using reported HPS cases and rodent reservoir abundance data.

Risk category	HPS Reported Cases		Rodent Reservoir Abundance	
	Risk Map	New cases	Risk Map	New cases
Low	606	21	208	6
Medium	21	4	273	12
High	13	2	143	14
Very High	5	6	21	1

DISCUSSION

This study is a first attempt to evaluate how different data predict disease risk, using the HPS disease as a model. Contrary to our predictions, both maps presented low values of similarities and correlation between them, predicting a large number of municipalities in different risk categories. In addition, RRA map performed better in predicting medium to high risk areas, while the RC risk map had greater accuracy in predicting low risk areas but performing worst in predict medium to high risk areas. Therefore, accuracy aspects varied between both risk maps, with this variation reflecting a non-uniform distribution of risk levels in São Paulo, predicted by the use of different data.

Reported HPS cases identified ~6% of the state of São Paulo (39 municipalities) as presenting medium-to-high risk of hantavirus transmission, while rodent reservoir abundance data (RRA) identified 68% (437

municipalities). Therefore, the RRA map had better sensitivity than the RC map, though the latter showed better specificity. It is noteworthy that the majority of municipalities are classified as medium to high risk in the RRA map, increasing the likelihood of hitting the new HPS cases. Withal, the RC map also classified the majority of the municipalities as low risk (96%), increasing the chances of hitting municipalities with no cases. Thus, both maps are overestimating one class of disease risk and failing to predict others classes.

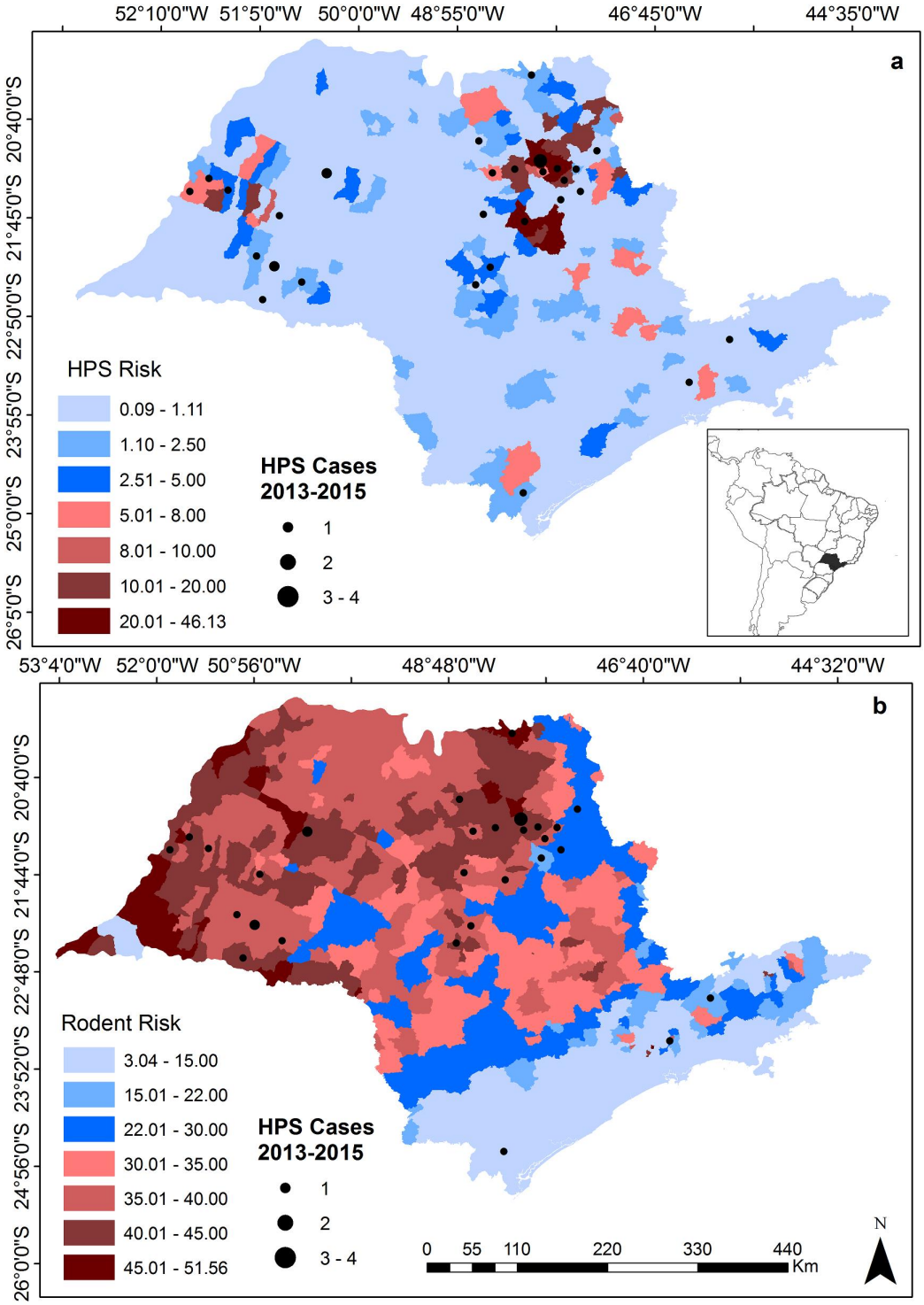

Figure 3. HPS risk maps based on (a) Reported cases for São Paulo State, Brazil, and (b) Rodent reservoir abundance, respectively. Municipalities with black outline are medium to very high HPS risk.

In our RRA map, we hypothesized that areas with high reservoir abundance were associated with a higher risk of infection, which is similar to what was explored by Ryan *et al.* (2004) and Diallo *et al.* (2011). Despite having high sensitivity and identifying a large number of new cases (82%), this map predicted a higher incidence of cases than actually occurred throughout the state. Generally, the emergence of human diseases is often found to be more spatially restricted than the distribution of the reservoir host (Schmaljohn & Hjelle 1997, Andreo *et al.* 2014), which seems to be the case for Hantavirus Pulmonary Syndrome in São Paulo. Therefore, it seems that our RRA map is strongly overestimating high risk, which may be due to the extrapolation of this risk map to the entire São Paulo State, using the original abundance data that was collected in Atlantic Forest areas. These differences highlight the level of uncertainty of our RRA model, which, together with the lack of data on cerrado areas, can be considered a source of bias; thus, more accurate measures of abundance that take into account typical cerrado landscapes could raise the performance of RRA maps. When comparing our rodent abundance extrapolations with data from published articles, we observe that extrapolated abundances of *N. lasiurus* differed from those studies by an average of 2.27 individuals. For *O. nigripes*, the mean difference was 3.36 individuals (Table 3). All studies but one differed from the extrapolated abundances by less than five individuals (N=11), but only four studies differed by less than one individual.

The RC map involves three interacting factors in space and time that are essential to a human get infected: an infected rodent; a certain abundance of reservoir rodents to proliferate the infection throughout the rodent population; and a susceptible human population. In other words, this proxy is based on a more complete data. However, the RC map had low sensitivity (36%) and high specificity (91%), better predicting the municipalities with zero cases. Therefore, it seems that this risk map is strongly overestimating low risk municipalities, which may be occurring because HPS is rare, with only a few reported cases (~200) in 20 years of data available.

Table 3. Comparison of rodent host abundances (*Necromys lasiurus*=NL and *Oligorizomys nigripes*=ON) between studies performed in São Paulo State and extrapolations by the RRA model. Differences between sampled and RRA model abundances smaller than one are in boldface. Sampling efforts are shown as trap-nights; Standardized abundance=(individuals/trap-night)*100; Sampling effort RRA=sampling effort of data used to extrapolate rodent abundance data.

Species	Municipality	Sampling effort (trap-nights)	Number of individuals captured	Standardized abundance	Sampling effort RRA	N° individuals predicted RRA	Standardized abundance predicted RRA	Difference between sampled and RRA model	Reference
NL	Pedreira	1,830	24	1.31	6,336	5.98	0.09	1.22	Bonvicino *et al.* 2002
NL	Sertãozinho	5,376	157	2.92	6,336	6.89	0.11	2.81	Gheler-Costa *et al.* 2013
NL	Angatuba	1,030	23	2.23	6,336	3.21	0.05	2.18	Martin *et al.* 2012
NL	Angatuba	893	36	4.03	6,336	3.87	0.06	3.97	Martin *et al.* 2012
NL	Angatuba	3,054	38	1.24	6,336	6.34	0.10	1.14	Martin *et al.* 2012
ON	**Pedreira**	**1,830**	**15**	**0.82**	**30,272**	**25.65**	**0.08**	**0.73**	**Bonvicino *et al.* 2002**
ON	**Rio Claro**	**911**	**5**	**0.55**	**30,272**	**50.54**	**0.17**	**0.38**	**Briani *et al* 2001**
ON	**Sertãozinho**	**5,376**	**29**	**0.54**	**30,272**	**20.20**	**0.07**	**0.47**	**Gheler-Costa *et al.* 2013**
ON	Cotia	2,700	34	1.26	30,272	1.10	0.00	1.26	Naxara *et al.* 2009
ON	Angatuba	1,030	122	11.84	30,272	29.30	0.10	11.75	Martin *et al.* 2012
ON	Angatuba	893	42	4.70	30,272	3.47	0.01	4.69	Martin *et al.* 2012
ON	Angatuba	3,054	130	4.26	30,272	2.20	0.01	4.25	Martin *et al.* 2012

As a first and novel approach, the analyses presented here were based on two types of easily accessible epidemiological data: official reported HPS cases and reservoir rodent abundances. Due to the limitations on the use of our RRA map, we suggest that data on viral infection and presence of infected rodents should be coupled with reservoir abundance data. The former may be the main determinant of direct transmission (Ostfeld *et al.* 2005), but it is a proxy that is hard to obtain or spatially extrapolate. The use of this data still needs to be tested and compared with host abundances and disease cases to assess whether it infers disease risk more accurately than the two used in this study. This is an opportune moment to evaluate the use and validity of different proxies, given the large number of emerging diseases affecting populations globally (Jones *et al*. 2008).

In conclusion. disease predictions using different proxies led to results that varied widely in their sensitivity and specificity in predicting new cases of HPS throughout São Paulo State. Each proxy had a bias and was effective in predicting only certain levels of disease risk. This represents a caveat to public policy makers, since similar results are normally used to guide the allocation of resources for preventive measures, educational campaigns and even the collection of hosts in the implementation of control policies. Therefore, the use of different proxies in modeling disease risk should be treated carefully and should be examined further in other studies, in other regions with different diseases and hosts, in order to determine the best data for inferring disease risk.

ACKNOWLEDGMENTS

São Paulo Research Foundation FAPESP (PP 2013/12515-5, RLM 2015/17739-4, JPM 2013/23457-6; MCR 2013/50421-2; RP 2005/56555-4), the Brazilian Research Council (MCR 312045/2013-1) and the National Council for Scientific and Technological Development (CNPQ, JPM307934/2011-0; AP 130997/2013-7) funded our study. R. Bovendorp gave useful suggestions on the manuscript. M. King, a native English speaker, revised the manuscript. We thank the anonymous referees for the invaluable comments.

REFERENCES

Andreo, V., Neteler, M., Rocchini, D., Provensal, C., Levis, S., Porcasi, X., Rizzoli, A., Lanfri, M., Scavuzzo, M., Pini, N., Enria, D., & Polop, J. 2014. Estimating hantavirus risk in Southern Argentina: A GIS-based approach combining human cases and host distribution. Viruses, 6(1), 201-222. DOI: 10.3390/v6010201

Batista, A. P. B., Mello, J. M., Raimundo, M. R., Scolforo, H. F., Reis, A. A., Scolforo, J. R. S. 2016. Species richness and diversity in shrub savana using ordinary kriging. *Pesquisa Agropecuária* Brasileira, 51(8), 951-966. DOI: 10.1590/S0100-204X2016000800008

Bhatt, S., Gething, P. W., Brady, O. J., Mesina, J. P., Farlow, A. W., Moyes, C. L., Drake, J. M., Brownstein, J. S., Hoen, A. G., Sankoh, O., Myers, M. F., George, D. F., Jaensich, T., Wint, G. R. W., Simmons, C. P., Scott, T. W., Farrar, J. J., & Hay, S. I. 2013. The global distribution and burden of dengue. Nature, 496(7446), 504-507. DOI:10.1038/nature12060

Biek, R., & Real, L. A. 2011. The landscape genetics of infectious disease emergence and spread. Molecular Ecology, 19(17), 3515-3531. DOI: 10.1111/j.1365-294X.2010.04679.x

Bonvicino, C. R., Lindbergh, S. M., & Maroja, L. S. 2002. Small non-flying mammals from conserved and altered areas of Atlantic Forest and Cerrado: comments on their potential use for monitoring environment. Brazilian Journal of Biology, 62(4B), 765-774. DOI: 10.1590/S1519-69842002000500005

Brazilian Ministry of Health. 2013 [Internet].Retrieved April 13, 2013, from http://portal.saude.gov.br/portal/saude/profissional/area.cfm?id_area=1558..

Briani, D. C., Santori, R. T., Vieira, M. V., & Gobbi, N. 2001. Mamíferos não-voadores de um fragmento de mata mesófila semidecídua do interior do Estado de São Paulo, Brasil. Holos Environment, Rio Claro, 1(2), 141-149.

Brooker, S., Hay, S. I., & Bundy, D. A. P. 2002. Tools from ecology: useful for evaluating infection risk models? Trends in Parasitology, 18(2), 70-74. DOI: 10.1016/S1471-4922(01)02223-1

CDC. 2014. HPS: Updated recommendations for risk reduction. CDC Publishing, Atlanta [Internet]. Retrieved from http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5109a1.htm

de Oliveira, S. V., Escobar, L. E., Peterson, A. T., & Gurgel-Gonçalves, R. 2013. Potential geographic distribution of hantavirus reservoirs in Brazil. PLoS ONE, 8(12): e85137. DOI: 10.1371/journal.pone.0085137

de Oliveira, S. V., Fonseca, L. X., Silva, P. M. R. B., Pereira, S. V. C., & Caldas, E. P. 2014. Análise do perfil epidemiológico da hantavirose no Brasil no período de 2007 a 2012. Journal of Tropical Pathology, 43(2), 131-142. DOI: 10.5216/rpt.v43i2.31115

de Oliveira, S. V., Fonseca, L. X., de Araújo Vilges, K. M., Maniglia, F. V. P., Pereira, S. V. C., de Caldas, E. P.,Tauil, P. L., & Gurgel-Gonçalves, R. 2015. Vulnerability of Brazilian municipalities to hantavirus infections based on multi-criteria decision analysis. Emerging Themes in Epidemiology, 12(1), 15. DOI: 10.1186/s12982-015-0036-5

Diallo, D., Talla, C., Ba, Y., Dia, I., Sall, A. A., & Diallo, M. 2011. Temporal distribution and spatial pattern of abundance of the Rift Valley fever and West Nile fever vectors in Barkedji, Senegal. Journal of Vector Ecology, 36(2), 426-436. DOI: 10.1111/j.1948-7134.2011.00184.x

Ferreira, M. S. 2003. Hantaviroses. Revista da sociedade brasileira de medicina tropical, 36(1), 81-96. DOI: 10.1590/S0037-86822003000100012

Gelman, A., & Parode, I. 2006. Bayesian measures of explained variance and pooling in multilevel (hierarchical) models. Technometrics, 48(2), 241-251. DOI: 10.1198/004017005000000517

Gheler-Costa, C., Sabino-Santos, Jr G., Amorim, L. S., Rosalino, L. M., Figueiredo, L. T. M., & Verdade, L. M. 2013. The effect of pre-harvest fire on the small mammal assemblage in sugarcane fields. Agriculture, Ecosystems & Environment, 171, 85-89. DOI: 10.1016/j.agee.2013.03.016

Gheler-Costa, C., Vettorazzi, C. A., Pardini, R., & Verdade, L. M. 2012. The distribution and abundance of small mammals in agroecosystems of Southeastern Brazil. Mammalia, 76(2), 185-191. DOI: 10.1515/mammalia-2011-0109

Glavanakov, S., White, D., Caraco, T., Lapenis, A., Robinson, G., Szymanski, B. K., & Maniatty, W. A. 2001. Lyme disease in New York state: spatial pattern at a regional scale. American Journal of Tropical Medicine and Hygiene, 65(5), 538-545.

Guerra, M., Walker, E., Jones, C., Paskewitz, S., & Cortinas, M. R. 2002. Predicting the risk of Lyme Disease: habitat suitability for Ixodes scapularis in the North Central United States. Emerging Infectious Diseases, 8(3), 289-297. DOI: 10.3201/eid0803.010166

Guis, H., Annelise Tran, Laroque, S. De, Baldet, T., Gerbier, G., Barragué, B., Biteau-Coroller, F., Roger, F., Viel, J. F., & Mauny, F. 2007. Use of high spatial resolution satellite imagery to characterize landscapes at risk for bluetongue. Veterinary Research, 38(5), 669-683. DOI: 10.1051/vetres:2007025

Guterres, A., de Oliveira, R. C., & Lemos, E. R. 2015. Is the evolution of Hantavirus driven by its host? Infection, Genetics and Evolution, 35,142-143. DOI: 10.1016/j.meegid.2015.08.002

Horobik, V., Keesing, F., & Ostfeld, R. S. 2007. Abundance and *Borrelia burgdorferi* infection prevalence of nymphal *Ixodes scapularis* ticks along forest - field edges. EcoHealth, 3(4), 262-268. DOI: 10.1007/s10393-006-0065-1

Kitron, U. 1998. Landscape ecology and epidemiology of vector-borne diseases: tools for spatial analysis. Journal of Medical Entomology, 35(4), 435-445. DOI: 10.1093/jmedent/35.4.435

Jackson, H. B., & Fahrig, L. 2012. What size is a biologically relevant landscape? Landscape Ecology, 27, 929-941.

Jones, K. E., Patel, N. G., Levy, M. A, Storeygard, A., Balk, D., Gittleman, J. L., & Daszak, P. 2008. Global trends in emerging infectious diseases. Nature, 451(7181), 990-993. DOI: 10.1038/nature06536

Jonsson, C. B., Figueiredo, L. T., & Vapalahti, O. 2010. A global perspective on hantavirus ecology, epidemiology, and disease. Clinical Microbiology Reviews, 23(2), 412-441. DOI: 10.1128/CMR.00062-09

Lee, H. W., French, G. R., & Lee, P. W. 1981. Observations on natural and laboratory infection of rodents with the etiologic agent of Korean hemorrhagic fever. The American Journal of Tropical Medicine and Hygiene, 30, 477-482.

Little, E., Barrera, R., Seto, K. C., & Diuk-wasser, M. 2011. Co-occurrence patterns of the dengue vector *Aedes aegypti* and *Aedes mediovitattus*, a dengue competent mosquito in Puerto Rico. EcoHealth, 8(3), 365-375. DOI: 10.1007/s10393-011-0708-8

Madhav, N. K., Brownstein, J. S., Tsao, J. I., & Fish, D. 2004. A dispersal model for the range expansion of Blacklegged Tick (Acari/ : Ixodidae). Journal of Medical Entomology, 41(5), 842-852. DOI: 10.1603/0022-2585-41.5.842

Martens, P., Kovats, R. S., Nijhof, S., Vries, P. De, Livermore, M. T. J., Bradley, D. J., Cox, J., & Mcmichael, A. J. 1999. Climate change and future populations at risk of malaria. Global Environmental Change, 9, 89-107. DOI: 10.1016/S0959-3780(99)00020-5

Martin, P. S., Gheler-Costa, C., Lopes, P. C., Rosalino, L. M., Verdade, L. M. 2012. Terrestrial non-volant small mammals in agro-silvicultural landscapes of Southern Brazil. Forest Ecology and Management, 282, 185-195. DOI: 10.1016/j.foreco.2012.07.002

McGarigal, K., Cushman, S. A., Neel, M. C., & Ene, E. 2002. Fragstats. Spatial Pattern Analysis Program for Categorical Maps. Computer software program produced by the authors at the University of Massachusetts, Amherst. Retrieved from http://www.umass.edu/landeco/research/fragstats/fragstats.html

Naxara, L., Pinotti B. T., & Pardini, R. 2009. Seasonal microhabitat selection by terrestrial rodents in an old-growth Atlantic Forest. Journal of Mammalogy, 90(2), 404-415. DOI: 10.1644/08-MAMM-A-100.1

Ostfeld, R. S., Glass, G. E., & Keesing, F. 2005. Spatial epidemiology: An emerging (or re-emerging) discipline. Trends in Ecology and Evolution, 20(6 SPEC. ISS.), 328-336. DOI: 10.1016/j.tree.2005.03.009

Pardini, R., Bueno, A. A., Gardner, T.A., Prado, P. I., & Metzger, J.P. 2010. Beyond the fragmentation threshold hypothesis: regime shifts in biodiversity across fragmented landscapes. PLoS ONE, 5(10), e13666. DOI: 10.1371/journal.pone.001366

Pereira, L. E., Suzuki, A., Bisord, I., Souza, R. P., Souza, L. T. M., Oshiro, F. M. Cerroni, M. P., Azevedo Neto, R. S., & Pinho, J. R. R. 2007. Estudo longitudinal da prevalência dos vírus Juquitiba e Araraquara em roedores das regiões da Mata Atlântica e do Cerrado do Brasil. Boletim Epidemiológico Paulista, 4(42), 2-13.

Prist, P. R., Uriarte, M., Tambosi, L. R., Prado, A., Pardini, R., DóAndrea, P. S., & Metzger, J. P. 2016. Landscape, environmental and social predictors of Hantavirus risk in São Paulo, Brazil. PLoS ONE, 11(10), e0163459. DOI: 10.1371/journal.pone.0163459

Reisen, W. K. 2010. Landscape epidemiology of vector-borne diseases. Annual Review of Entomology, 55(1), 451-483. DOI: 10.1146/annurev-ento-112408-085419

Ribeiro, M. C., Metzger, J. P., Martensen, A. C., Ponzoni, F. J., & Hirota, M. M. 2009. The Brazilian Atlantic Forest: how much is left, and how is the remaining forest distributed? Implications for conservation. Biological Conservation, 142(6), 1141-1153. DOI: 10.1016/j.biocon.2009.02.021

Rogers, D. J., & Randolph, S. E. 2000. The global spread of Malaria in a future, warmer world. Science, 289(5485), 1763. DOI: 10.1126/science.289.5485.1763

Ryan, P. A., Lyons, S. A., Alsemgeest, D., Thomas, P., & Kay, B. H. 2004. Spatial statistical analysis of adult mosquito (Diptera: Culicidae) counts: an example using light trap data, in Redland Shire, Southeastern Queensland, Australia. Journal of Medical Entomology, (Taylor 1984),41(6), 1143-1156. DOI: 10.1603/0022-2585-41.6.1143

Schmaljohn, C., & Hjelle, B. 1997. Hantaviruses: A global disease problem. Emerging Infectious Diseases, 3(2), 95-104. DOI: 10.3201/eid0302.970202

Tälleklint, L., Jaenson, T. G. T., & Jaenson, G. T. 1995. Is the small mammal (*Clethrionomys glareolus*) or the tick vector (*Ixodes ricinus*) the primary overwintering reservoir for the Lyme Borreliosis Spirochete in Sweden? Journal of Wildlife Diseases, 31(4), 537-540. DOI: 10.7589/0090-3558-31.4.537

Umetsu, F. 2010. Efeito do contexto da paisagem em diferentes escalas sobre a distribuição de pequenos mamíferos em cultivos agrícolas e em remanescentes florestais. In: Instituto de Biociência. Departamento de Zoologia. Universidade de São Paulo, São Paulo, p. 88.

Umetsu, F., & Pardini, R. 2007. Small mammals in a mosaic of forest remnants and anthropogenic habitats - evaluating matrix quality in an Atlantic Forest landscape. Landscape Ecology, 22(4), 517-530. DOI:10.1007/s10980-006-9041-y

Vapalahti, K., Virtala, A. M., Vaheri, A., & Vapalahti, O. 2010. Case-control study on Puumala virus infection: smoking is a risk factor. Epidemiology & Infectiontion, 138(4), 576-584. DOI: 10.1017/S095026880999077X

Willemann, M. C. A. & de Oliveira, S. V. 2014. Risk factors associated with hantavirosis fatality: a regional analysis from a case-control study in Brazil. Revista da Sociedade Brasileira de Medicina Tropical, 47(1), 47-51. DOI: 10.1590/0037-8682-0243-2013.

Xiao, H., Huang, R, Gao, L. D., Huang, C. R., Lin, X. L., Li, N., Liu, H. N., Tong, S. L., & Tiam, H. Y. 2016. Effects of humidity variation on the Hantavirus infection and Hemorrhagic Fever with Renal Syndrome occurrence in Subtropical China. American Journal of Tropical Medicine and Hygiene, 94(2), 420-427. DOI:10.4269/ajtmh.15-0486

Submitted: May 15 2016.
Accepted: December 15 2016.